

McKinzey, R.K., & Campagna, V. (April 30, 2002). Authors' reply to prior review of "Rorschach Interscorer agreement". *WebPsychEmpiricist*. Retrieved (date) from: [http://www.wpe.info/papers\\_table](http://www.wpe.info/papers_table)

# WPE WebPsychEmpiricist

## Authors' Reply to Prior Review of "Rorschach Interscorer Agreement"

R. K. McKinzey

Oakland, CA

Victoria Campagna

Redwood City, CA

### Abstract

The authors of "The Rorschach, Exner's Comprehensive System, Interscorer Agreement, and Death" included six blind, anonymous reviews offered during previous submissions. The paper's authors reply, arguing that the paper was wrongly rejected, the sole conclusion remains valid, and another study has already partially replicated the results.

### Authors' Reply to Prior Review: "Interscorer Agreement"

The reviewers wrongly rejected this paper. They argued that this paper should not be published because there were so many alternative explanations for the poor agreement other than the inadequate nature of the CS. These included poor protocol administration, demand characteristics, patient characteristics, and unstandardized format.

The reviewers all miss the point. We acknowledge there are alternative explanations, identify several, and discuss them. However, none of these explanations refute our single conclusion: Even using the stringent, elaborate rules provided by the CS, the Rorschach is easily rebutted, and for many reasons.

The literature now refutes some of the alternative explanations. The notion that the putatively poor protocol administration or unstandardized format can account for the dramatic levels of disagreement is not supported. In our last paragraph, we argued that the next step was to repeat the experiment, this time with an exemplary protocol. While this paper was under review, such an experiment was published (Guarnaccia, Dill, Sabatino, & Southwick, 2001). Using authoritative textbook responses of both patients and nonpatients, they constructed a protocol whose administration, formatting, inquiry, and scoring were impeccable. They then sent it to 33 independent scorers, 20 students and 12 professionals. Many errors were made by both students and professionals, and most were made on patient responses. Of the 10 scoring categories done by professionals on patient responses (the closest to this study's design), 6 (60%) failed to reach acceptable accuracy levels. It would seem that the rationale used by JCP and PPRP's editors to reject our paper lacks empirical support.

Other alternate explanations fail on logical grounds. Blaming poor reliability on patient characteristics (Reviewer B) is a form of Special Pleading. How can a test of psychopathology be defended if the very people it is designed to test produce unreliable results?

As we noted in the Footnote 12, one of our participants preceded Reviewer 3 in blaming the poor reliability on demand characteristics. That is, they argued that the 90% of psychologists that offered scores might not have done so without the structure of the research. Such an argument ignores the demand characteristics of forensic rebuttal: If there is a way of producing a different score (and thus a different opinion), it will be found.

Reviewers 3 and C question the generalizability of the study. Can anything be learned from the one protocol to many raters design? It would seem so, in the opinion of our and

Guarnaccia et al's (2001) 65 participants. It seems that errors and disagreements rise dramatically when many independent raters score the CS, echoing the more modest disagreements found when two less independent raters score many protocols.

Reviewers 1 and B argue that no single data point could possibly be used to make a forensic point or diagnosis. This argument is disingenuous, used whenever a test, scale, score, or decision rule fails cross-validation. In this case, it ignores the fact that Dr. Doe did use two constellations of one test, without supporting history, to make a forensic diagnosis.

In summarizing our study's design, Reviewer 1 comments that, "Raters were polled concerning the codes and quality of administration." This is incorrect. We did not solicit any participant's comments. They were offered freely. We thought they were of import, so we reported them. Similarly, Reviewer 3 misunderstood—we did not ask any one to interpret this protocol, just to score it. The study is about scoring, not interpretation.

Finally, Reviewer C warned that permissions should be obtained from Dr. Doe and the patient. No such permissions are needed, since, as we noted in the Materials section, the protocol and scorings are part of public record.

After the rejection by JCP, we used a new feature of APA's 5<sup>th</sup> style manual (American Psychological Association, 2001). This feature allows contributors to warn the editor of possible bias in the pool of available reviewers (pages 303-304 & 411). Consequently, we wrote in our cover letter:

As APA 5 requires us to state, we have no financial interests in the Rorschach industry, such as research grants, teaching courses, conducting seminars, authoring books, or having an ongoing forensic case using the Rorschach. We assume your reviewers will not have any such interests either. Given the findings, the eminent scholars that are part of the Rorschach industry will undoubtedly want to provide commentary.

After getting the PPRP reviews, we asked Dr. Kenkle if our request had been honored. She replied (3/29/02):

I fully agree with your aim to have a fair and objective peer review process. In selecting reviewers for any ms, I and the associate editors of *PP* try to ensure a balanced representation. If an issue is controversial, we insure that the panel is not weighted in either a pro or con direction. Depending on the topic, we seek balance in terms of career setting (academic and practice), theoretical orientation, gender or racial/ethnic composition, etc., etc.

However, we do not have or collect the type of information you have outlined on reviewers. That is information we systematically collect on each author of a submitted

manuscript, but not on reviewers. To do so would be unwieldy and, I believe, unnecessary since it is a key responsibility of the editors to insure the fairness of the review process.

I can assure you that in putting together the list of reviewers for your manuscript, that I was conscientious about creating a balanced panel. Given the differences in reviewers' opinions about the Exner system and/or the Rorschach, there still was great consistency in the feedback on the manuscript. Therefore, it seems likely that the reviews were based on a high level of objectivity and an evaluation of the research methodology.

For reasons now obvious, we disagree. We recommend editors handling research critical of the CS pay particular attention to possible reviewer bias. People whose income relies on the reputation of the CS should not be in the position of anonymously suppressing research.

#### References

- American Psychological Association. (2001). *Publication Manual of the American Psychological Association* ( 5 ed.).
- Guarnaccia, V., Dill, C. A., Sabatino, S., & Southwick, S. (2001). Scoring accuracy using the Comprehensive System for the Rorschach. *Journal of Personality Assessment*, 77(3), 464-474.